ELSEVIER Linear Algebra and its Applications 278 (1998) 63–84

LINEAR ALGEBRA AND ITS APPLICATIONS

Schur complements obey Lambek's categorial grammar: Another view of Gaussian elimination and LU decomposition

D. Stott Parker [1]

Computer Science Department, University of California, Los Angeles, CA 90095-1596, USA

Received 6 September 1995; accepted 14 October 1997

Submitted by R.A. Brualdi

Abstract

For three decades Schur complements have seen increasing applications in linear algebra, often as abstractions of Gaussian elimination. It is known that they obey certain nontrivial identities, such as Crabtree and Haynsworth's *quotient property*. We began this work asking if there were a theory for deciding their properties in general.

Lambek's Categorial Grammar is a deductive system formalized in 1958 by Lambek as a mathematical foundation for a syntactic calculus of language. We show that Categorial Grammar gives a deductive system for deriving identities obeyed by LU- and UL-decompositions, Gaussian elimination, and Schur complements.

At first impression this seems to be a strange result, connecting two unrelated topics. In retrospect, though, it is a consequence of the way both use quotients. It may have applications in developing grammatical formalisms and numerical algorithms. © 1998 Elsevier Science Inc. All rights reserved.

Keywords: Schur complements; Quotients; Categorial grammar; Categorical grammar; Lambek calculus; Gaussian elimination; LU decomposition

1. Introduction

Although Schur complements are mentioned in almost every modern text on computational linear algebra, and have deep connections to the theory of

[1] E-mail: stott@cs.ucla.edu.

PII: S0024-3795(97)10033-7

quotients over modules and homological algebra, their literature is less than 30 years old and still very incomplete. For example, Crabtree and Haynsworth [7] first noted the 'quotient identity' for Schur complements (discussed below) in 1969, and there have been many papers relating them to generalized inverses. Beyond Cottle's 1974 survey [6] we have found no general characterization of identities obeyed by Schur complements.

Categorial Grammar is a simple yet powerful kind of formal grammar, in ways akin to Chomsky's Context Free Grammar, that has been inspirational in computational linguistics. It developed from work of Ajdukiewicz published in 1935 [1], with improvements by Bar-Hillel in 1953 [2], ch. 5. (In 1951 Evans independently developed a similar axiom system, without associativity, for word problems, loops, and groupoids [8].) In 1958 Lambek [15] extended Categorial Grammar into a complete formal system (a syntactic calculus with a decision procedure), an outcome of his "observation that a notation which was useful in two branches of algebra, module theory and ideal theory..., could also be applied to the study of sentence structure in natural languages" [17]. Good historical summaries of work on Categorial Grammar, Lambek calculi, and related formal systems are Bar-Hillel's book [2], chs. 6, 8, the edited volume [21], and the recent survey by Moortgat [20].

In studying Schur complements and Gaussian elimination [24–26], we noticed that matrix decompositions using Schur complements (and thus the decompositions obtained with Gaussian elimination) obey the rules of Categorial Grammar. The purpose of this paper is merely to sketch various ways of applying Categorial Grammar, and to illustrate their potential in reasoning about Schur complements.

2. Schur complements

2.1. Terminology

Definition 1. The *submatrix of A with rows $i_1, \ldots, i_p$ and columns $j_1, \ldots, j_q$* (in that order) is denoted $A[i_1, \ldots, i_p | j_1, \ldots, j_q]$, and consists of the indicated elements of A:

$$A[i_1, \ldots, i_p | j_1, \ldots, j_q] = \begin{pmatrix} a_{i_1,j_1} & a_{i_1,j_2} & \cdots & a_{i_1,j_q} \\ a_{i_2,j_1} & a_{i_2,j_2} & \cdots & a_{i_2,j_q} \\ \vdots & \vdots & \ddots & \vdots \\ a_{i_p,j_1} & a_{i_p,j_2} & \cdots & a_{i_p,j_q} \end{pmatrix}.$$

For example, the submatrix of A with rows 3, 2 and columns 1, 3, 1, 4 is

$$A[3,2\,|\,1,3,1,4] = \begin{pmatrix} a_{31} & a_{33} & a_{31} & a_{34} \\ a_{21} & a_{23} & a_{21} & a_{24} \end{pmatrix}.$$

Also, when $p = 0$ or $q = 0$ the submatrix is 1 by convention, so $A[\,|\,] = 1$. This convention is natural here, since for matrix direct sums the empty matrix (the 0×0 matrix) is the identity.

For any subsequence α of $1, \ldots, n$, $A[\alpha|\alpha]$ is a *principal submatrix* of the $n \times n$ matrix A.

For $1 \leqslant k \leqslant n$, $A[1, \ldots, k|1, \ldots, k]$ is a *leading principal submatrix* of A.

For $1 \leqslant k \leqslant n$, $A[k, \ldots, n|k, \ldots, n]$ is a *trailing principal submatrix* of A.

Definition 2. If X and Y are square matrices, define

$Y \trianglelefteq X$ iff Y is a nonsingular leading principal submatrix of X.

$X \trianglerighteq Y$ iff Y is a nonsingular trailing principal submatrix of X.

Definition 3.

$$X \oplus Y = \begin{pmatrix} X & 0 \\ 0 & Y \end{pmatrix}$$

is the *matrix direct sum* of square matrices X and Y.

Thus $1 \oplus X = X \oplus 1 = X$. Also, if Y is nonsingular, $Y \trianglelefteq (Y \oplus X)$ and $(X \oplus Y) \trianglerighteq Y$.

Definition 4. An $n \times n$ matrix A is *nondegenerate* if all principal submatrices are nonsingular, i.e., $\det A[\alpha|\alpha] \neq 0$, for every subsequence α of $1, \ldots, n$.

Throughout this paper we will require all matrices to be nondegenerate. This requirement is not as restrictive as it may seem, since any nonsingular linear system $Ax = b$ can be transformed to an equivalent nondegenerate system $\tilde{A}x = \tilde{b}$ via appropriate random matrices [24,26].

2.2. The Schur complement and its properties

Emilie Haynsworth pioneered interest in the Schur complement in a series of papers published between 1968 and 1974 [5,7,10–12,22]. Cottle [6] surveys various places where Schur complements turn up, including determinantal

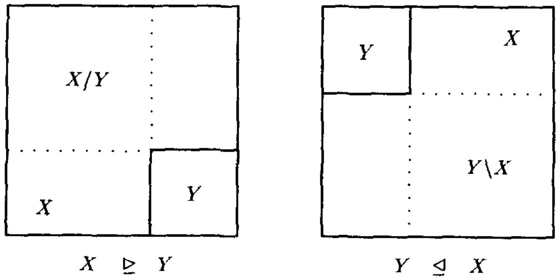


Fig. 1. The Schur complement of a trailing or leading principal submatrix Y in X.

identities, pivoting, inverse matrices, inertia, etc. See also [13], pp. 17–23. Here we use a definition in which the ordering of rows and columns is significant.

Definition 5. If a nondegenerate matrix X has block decomposition

$$X = \begin{pmatrix} A & B \\ C & D \end{pmatrix},$$

so that $A \trianglelefteq X$ and $X \trianglerighteq D$ then:
the *trailing Schur complement of D in X*, written (X/D), is

$$(X/D) = A - BD^{-1}C,$$

and the *leading Schur complement of A in X*, written $(A \setminus X)$, is

$$(A \setminus X) = D - CA^{-1}B.$$

Schur complements can be understood pictorially. In Fig. 1, the dotted outlines in the diagrams delimit the complement of the submatrix Y in X, if for example the off-diagonal blocks are zero. (Nonvisually oriented readers can ignore these figures without really losing anything, but they are handy on occasion, especially in reasoning about block matrix decompositions later.)

Lemma 1 (Schur's identity for determinants [27]). *If*

$$X = \begin{pmatrix} A & B \\ C & D \end{pmatrix},$$

so that $A \trianglelefteq X$ and $X \trianglerighteq D$, then

$$\det X = \det(X/D) \quad \det D,$$
$$\det X = \det A \quad \det(A \setminus X).$$

These identities are readily derivable with a block LU decomposition of X.

Lemma 2 (Quotient property of Schur complements). *If X is a nondegenerate matrix such that*

$$X = \begin{pmatrix} A & B \\ C & D \end{pmatrix}, \quad A = \begin{pmatrix} A_{11} & A_{12} \\ A_{21} & A_{22} \end{pmatrix}, \quad D = \begin{pmatrix} D_{11} & D_{12} \\ D_{21} & D_{22} \end{pmatrix},$$

are block decompositions with $A_{11} \trianglelefteq A \trianglelefteq X$ and $X \trianglerighteq D \trianglerighteq D_{22}$, then

$$\begin{aligned}(X/D) &= ((X/D_{22})/(D/D_{22})),\\ (A \setminus X) &= ((A_{11} \setminus A) \setminus (A_{11} \setminus X)).\end{aligned}$$

This was proven in [7], using Schur's identity for determinants (Lemma 1). The proof is complicated. Another proof by Ostrowski [22] is also complicated. Brualdi and Schneider [4] give two simple proofs based on determinantal identities and Gaussian elimination.

2.3. Gaussian elimination and Schur complements

In its basic form, Gaussian elimination is a sequence of transformations to an $n \times n$ square matrix $A = (a_{ij})$, reducing it to upper-diagonal form in n steps. It can be defined equationally, with the initial assignment $a_{ij}^{(1)} = a_{ij}$ and the recursive definition of $a_{ij}^{(k+1)}$ for $1 \leqslant k \leqslant n-1$:

$$a_{ij}^{(k+1)} = \begin{cases} 0 & i \geqslant k+1, j = k, \\ a_{ij}^{(k)} - a_{ik}^{(k)}\left(a_{kk}^{(k)}\right)^{-1} a_{kj}^{(k)} & i \geqslant k+1, j \geqslant k+1, \\ a_{ij}^{(k)} & \text{otherwise.} \end{cases}$$

Often this definition is viewed as summarizing a program that implements Gaussian elimination. However, in this paper, the definition is viewed as a set of equations, with no 'roundoff error'. Gaussian elimination is typically implemented with for-loops in a program requiring $n(n^2-1)/3$ assignments altogether:

```
for k = 1 to n − 1 do
  for i = k + 1 to n do
    begin
      m_ik := a_ik ÷ a_kk;                 { a_ik^(k) ÷ a_kk^(k) }
      for j = k + 1 to n do
          a_ij := a_ij − m_ik × a_kj       { a_ij^(k) − a_ik^(k) ÷ a_kk^(k) × a_kj^(k) }
    end
```

Starting with $A^{(1)} = A$, iterations of the outer loop compute $A^{(k+1)} = (a_{ij}^{(k+1)})$ for $1 \leqslant k \leqslant n-1$, except that as written above the program does not zero the elements of A below the diagonal in the first k columns. (These elements are in precisely the same positions as the multipliers m_{ik}, so many implementations store m_{ik} in a_{ik}.) With the program we can form the triangular matrices

$$L = \begin{pmatrix} 1 & & & & \\ m_{21} & 1 & & & \\ m_{31} & m_{32} & 1 & & \\ \vdots & \vdots & \vdots & \ddots & \\ m_{n1} & m_{n2} & m_{n3} & \cdots & 1 \end{pmatrix}, \quad U = \begin{pmatrix} a_{11}^{(n)} & a_{12}^{(n)} & a_{13}^{(n)} & \cdots & a_{1n}^{(n)} \\ & a_{22}^{(n)} & a_{23}^{(n)} & \cdots & a_{2n}^{(n)} \\ & & a_{33}^{(n)} & \cdots & a_{3n}^{(n)} \\ & & & \ddots & \vdots \\ & & & & a_{nn}^{(n)} \end{pmatrix}$$

such that $A^{(n)} = U$ and $A = LU$. Gaussian elimination works only when the $a_{kk}^{(k)}$ are nonzero, for $1 \leqslant k \leqslant n$, which holds when A is nondegenerate [25]. In this paper we do not consider pivoting, or other extensions of Gaussian elimination for degenerate matrices.

Definition 6. Let $A_{(1)} = A$ and, for $1 \leqslant k < n$,

$$A_{(k+1)} \stackrel{\text{def}}{=} A^{(k+1)}[(k+1),\dots,n|(k+1),\dots,n],$$

i.e., $A_{(k+1)}$ is the $(n-k) \times (n-k)$ submatrix $(a_{ij}^{(k+1)})$ where $1 \leqslant k < i,j \leqslant n$.

Theorem 1. *If A is an $n \times n$ nondegenerate matrix, then*

$$A_{(k+1)} = (A[1,\dots,k\,|\,1,\dots,k] \setminus A).$$

Equivalently: $a_{ij}^{(k+1)} = (A[1,\dots,k\,|\,1,\dots,k] \setminus A\ [1,\dots,k,i\,|\,1,\dots,k,j])$.

Proved in [9,25]. Thus every result of Gaussian elimination is expressible as a Schur complement.

Corollary 1. *If A is an $n \times n$ nondegenerate matrix, then*

$$A_{(k+1)} = ((a_{kk}^{(k)}) \setminus A_{(k)}), \quad 1 \leqslant k < n.$$

This identity combines Lemma 2 (the quotient property) and Theorem 1. It gives an incremental algorithm for computing $A_{(k+1)}$:

$A_{(1)} := A$;
for $k = 1$ **to** $n-1$ **do**
  begin
    $a_{kk}^{(k)} := A_{(k)}[1|1]$;
    $A_{(k+1)} := ((a_{kk}^{(k)}) \setminus A_{(k)})$
  end.

This is exactly Gaussian elimination.

3. Lambek's Categorial Grammar

In Categorial Grammar, words are assigned "categories", which are intended as "types", in the philosophical sense attributed to Aristotle and Kant. We can introduce a *lexicon* – a binary relation between words and types – such as the following:

Word	Type
equations	N
Gauss	NP
linear	N/N
quickly	S\S
solved	NP\(S/NP)
slept	NP\S
the	NP/N

Each type is a functional *pattern*, describing the role or effect of the word in a sentence. The word "Gauss" is a Noun Phrase. Verbs such as "slept" yield sentences when composed with noun phrases, so intuitively they have a type like "NP $\mapsto$ S". In Categorial Grammar this type is either NP\S or S/NP, which also captures insistence on word orderings. Specifically, NP\S maps an *immediately preceding* NP to an S, while S/NP would map an *immediately following* NP to an S, so the word "slept" of type NP\S yields a sentence if there is a preceding noun phrase, but not when there is a following noun phrase.

The preceding/following ordering relation is the precedence relation among elements in sequences, and here sequences are constructed using the associative '$\cdot$' operator. As an example, (S/NP)$\cdot$ NP$\cdot$(S\S) is such a sequence, and in it S/NP has a following NP, and NP has a following S\S.

Generally a *type* is defined to be either an *atomic type* (a symbol like NP or S) or a *compound type* (an expression made up of types and the binary operators '$\cdot$', '\', and '/'). Natural language categories (verb, adjective, etc.) are typically identifiable with a small number of types, and a particular word can therefore be assigned more than one type. For example, adverbs like "quickly" could be identified with either S\S or S/S, but for simplicity only one type is included in the lexicon above.

Types obey certain laws. For example we would expect adverbs, of type S\S, also to be of type (NP\S)\(NP\S), mapping verbs to verbs. Categorial Grammar defines reduction laws with which these intuitive subsumptions among types can be derived formally, and with which types can be reduced (simplified, or parsed). In [17], Lambek gives axioms and rules of inference defining the relation '$\Rightarrow$' on types X, Y, Z where 1 is a special identity type:

Axioms	$(X \cdot 1) \Rightarrow X$	
	$X \Rightarrow (X \cdot 1)$	
	$(1 \cdot X) \Rightarrow X$	
	$X \Rightarrow (1 \cdot X)$	
	$((X \cdot Y) \cdot Z) \Rightarrow (X \cdot (Y \cdot Z))$	
	$(X \cdot (Y \cdot Z)) \Rightarrow ((X \cdot Y) \cdot Z)$	
Rules of inference	if $X \Rightarrow Y$ and $Y \Rightarrow Z$	then $X \Rightarrow Z$
	if $(Y \cdot Z) \Rightarrow X$	then $Y \Rightarrow (X/Z)$
	if $Y \Rightarrow (X/Z)$	then $(Y \cdot Z) \Rightarrow X$
	if $(Y \cdot Z) \Rightarrow X$	then $Z \Rightarrow (Y \setminus X)$
	if $Z \Rightarrow (Y \setminus X)$	then $(Y \cdot Z) \Rightarrow X$

From these rules '$\Rightarrow$' can be seen to be reflexive (the first two axioms and first inference rule give $X \Rightarrow X$), and transitive (first inference rule). Categorial Grammar is sometimes defined differently, but this definition is suitable for our needs.

With this formal system we can derive many interesting laws. For example, using A and B as type variables instead of X and Y to avoid confusion, we can derive the so-called 'type raising' law

$$A \;\Rightarrow\; ((B/A) \setminus B)$$

in the following way:

1. $B/A \;\Rightarrow\; ((B/A) \cdot 1)$	axiom:	$X \Rightarrow (X \cdot 1)$
2. $((B/A) \cdot 1) \;\Rightarrow\; B/A$	axiom:	$(X \cdot 1) \Rightarrow X$
3. $B/A \;\Rightarrow\; B/A$	1. & 2. & rule:	if $X \Rightarrow Y$ and $Y \Rightarrow Z$ then $X \Rightarrow Z$
4. $((B/A) \cdot A) \Rightarrow B$	3. & rule:	if $Y \Rightarrow (X/Z)$ then $(Y \cdot Z) \Rightarrow X$
5. $A \Rightarrow ((B/A) \setminus B)$	4. & rule:	if $(Y \cdot Z) \Rightarrow X$ then $Z \Rightarrow (Y \setminus X)$.

In the same way we can derive each of the following reduction laws [18], p. 11ff:

Application	$(X/Y) \cdot Y$	$\Rightarrow$	X
	$Y \cdot (Y \setminus X)$	$\Rightarrow$	X
Composition	$(X/Y) \cdot (Y/Z)$	$\Rightarrow$	(X/Z)
	$(Z \setminus Y) \cdot (Y \setminus X)$	$\Rightarrow$	$(Z \setminus X)$
Associativity	$(Z \setminus X)/Y$	$\Rightarrow$	$Z \setminus (X/Y)$
	$Z \setminus (X/Y)$	$\Rightarrow$	$(Z \setminus X)/Y$
Type raising	X	$\Rightarrow$	$Y/(X \setminus Y)$
	X	$\Rightarrow$	$(Y/X) \setminus Y$
Quotient	(X/Y)	$\Rightarrow$	$(X/Z)/(Y/Z)$
	(Y/Z)	$\Rightarrow$	$(X/Y) \setminus (X/Z)$
	$(Y \setminus X)$	$\Rightarrow$	$(Z \setminus Y) \setminus (Z \setminus X)$
	$(Z \setminus Y)$	$\Rightarrow$	$(Z \setminus X)/(Y \setminus X)$.

For example, the four quotient laws are derivable directly from the composition laws with the rules of inference.

In addition, we can derive rules of inference, including [17], p. 303:

if $Y \Rightarrow Z$ then $(X \cdot Y) \Rightarrow (X \cdot Z)$
if $Y \Rightarrow Z$ then $(Y \cdot X) \Rightarrow (Z \cdot X)$
if $Y \Rightarrow Z$ then $(X/Z) \Rightarrow (X/Y)$
if $Y \Rightarrow Z$ then $(Y/X) \Rightarrow (Z/X)$
if $Y \Rightarrow Z$ then $(Z \setminus X) \Rightarrow (Y \setminus X)$
if $Y \Rightarrow Z$ then $(X \setminus Y) \Rightarrow (X \setminus Z)$.

It is convenient to use these derived laws and rules of inference in order to parse a sequence of words by reducing it to the type 'sentence' (represented above by 'S'). The following example shows how this can work, repeatedly using the derived rules of inference, the indicated derived laws (on the underlined subexpression), and transitivity:

	Gauss	·	solved	·	the	·	linear	·	equations	·	quickly
	NP	·	NP\(S/NP)	·	NP/N	·	N/N	·	N	·	S\S
$\Rightarrow$Composition	NP	·	NP\(S/NP)	·	NP/N			·	N	·	S\S
$\Rightarrow$Application	NP	·	NP\(S/NP)	·	NP					·	S\S
$\Rightarrow$Associativity	NP	·	(NP\S)/NP	·	NP					·	S\S
$\Rightarrow$Application	NP	·	(NP\S)							·	S\S
$\Rightarrow$Quotient	NP	·	(NP\S)							·	(NP\S)\(NP\S)
$\Rightarrow$Application	NP	·	NP\S								
$\Rightarrow$Application	S										

In other words, this derivation gives the sketch of a proof that

$$\mathsf{NP} \cdot \mathsf{NP} \setminus (\mathsf{S}/\mathsf{NP}) \cdot \mathsf{NP}/\mathsf{N} \cdot \mathsf{N}/\mathsf{N} \cdot \mathsf{N} \cdot \mathsf{S} \setminus \mathsf{S} \Rightarrow \mathsf{S}.$$

A *categorial grammar* for a language is the syntactic calculus above, along with a lexicon that assigns types to individual words in the language. With this grammar we can introduce a suitable definition for sentences in terms of types. For example, usually sequences of words $(w_1 \cdot w_2 \cdots w_n)$ whose lexical types $X_1, \ldots, X_n$ satisfy $(X_1 \cdot X_2 \cdots X_n) \Rightarrow S$ are called *sentences* of the language.

By now the reader will have gotten the drift here – clearly there is a strong parallel between Schur complements and Lambek's Categorial Grammar. The question is how to formalize the connection.

One answer is simply to ignore the matrices per se, and focus only on submatrix indices, which are sequences like the sequences of words above: e.g., $A[1\,|\,1] = (A[1,2,3\,|\,1,2,3] \,/\, A[2,3\,|\,2,3])$. This is a direct connection, and we will return to it in Section 8.

Another answer is to view an individual matrix as generating a set of its 'block unit diagonalizations', which are block diagonal matrix decompositions using Schur complements. These sets have a natural quotient structure like that behind Categorial Grammar.

4. ULU decompositions and ULU classes

Definition 7. An $n \times n$ block unit lower triangular matrix L has form

$$\begin{pmatrix} I_m & 0 \\ M & I_{m-n} \end{pmatrix}$$

for some $1 \leqslant m < n$, where I_m is an $m \times m$ identity matrix, and M is an $(n-m) \times m$ arbitrary matrix. Similarly, a block unit upper triangular matrix U has form

$$\begin{pmatrix} I_{m-n} & M \\ 0 & I_m \end{pmatrix}.$$

The block matrix

$$X = \begin{pmatrix} A & B \\ C & D \end{pmatrix}$$

has the block unit UL diagonalization $(X/D) \oplus D$, and the block unit LU diagonalization $A \oplus (A \setminus X)$, provided A and D are square and nonsingular.

These definitions reflect the following decompositions:

$$\begin{aligned}
\begin{pmatrix} I & -BD^{-1} \\ 0 & I \end{pmatrix}\begin{pmatrix} A & B \\ C & D \end{pmatrix}\begin{pmatrix} I & 0 \\ -D^{-1}C & I \end{pmatrix} &= \begin{pmatrix} (A - BD^{-1}C) & 0 \\ 0 & D \end{pmatrix} \\
&= (X / D) \oplus D, \\
\begin{pmatrix} I & 0 \\ -CA^{-1} & I \end{pmatrix}\begin{pmatrix} A & B \\ C & D \end{pmatrix}\begin{pmatrix} I & -A^{-1}B \\ 0 & I \end{pmatrix} &= \begin{pmatrix} A & 0 \\ 0 & (D - CA^{-1}B) \end{pmatrix} \\
&= A \oplus (A \setminus X).
\end{aligned}$$

For example, the matrix

$$X = \begin{pmatrix} 2 & 3 & 5 \\ 3 & 5 & 7 \\ 1 & 2 & 3 \end{pmatrix},$$

with determinant 1 has the following block unit diagonalizations:

$$
\begin{aligned}
(X \,/\, X[3\,|\,3]) \oplus X[3\,|\,3] &= \begin{pmatrix} 1/3 & -1/3 & \\ 2/3 & 1/3 & \\ & & 3 \end{pmatrix} \\
(X \,/\, X[2,3\,|\,2,3]) \oplus X[2,3\,|\,2,3] &= \begin{pmatrix} 1 & & \\ & 5 & 7 \\ & 2 & 3 \end{pmatrix} \\
X[1\,|\,1] \oplus (X[1\,|\,1] \setminus X) &= \begin{pmatrix} 2 & & \\ & 1/2 & -1/2 \\ & 1/2 & 1/2 \end{pmatrix} \\
X[1,2\,|\,1,2] \oplus (X[1,2\,|\,1,2] \setminus X &= \begin{pmatrix} 2 & 3 & \\ 3 & 5 & \\ & & 1 \end{pmatrix}.
\end{aligned}
$$

This diagonalization process can be repeated on submatrices, until the results are diagonal. Fig. 2 diagrams the set of block unit diagonalizations for the matrix X, showing how they are derived. Notice each member of the set has the same determinant. Also the diagonal matrix D in the LDU decomposition of X is the fifth matrix on the bottom row.

These recursive block unit diagonalizations ultimately yield a (recursive) block LU- or UL-decomposition. Although traditionally one uses either LU- or UL-decompositions exclusively, here they can be interleaved. For example:

$$
\begin{aligned}
\begin{pmatrix} 2 & 3 & 5 \\ 3 & 5 & 7 \\ 1 & 2 & 3 \end{pmatrix} &\overset{\text{LU}}{=} \begin{pmatrix} 1 & & \\ & 1 & \\ -1 & 1 & 1 \end{pmatrix} \begin{pmatrix} 2 & 3 & \\ 3 & 5 & \\ & & 1 \end{pmatrix} \begin{pmatrix} 1 & & 4 \\ & 1 & -1 \\ & & 1 \end{pmatrix} \\
&\overset{\text{UL}}{=} \begin{pmatrix} 1 & & \\ & 1 & \\ -1 & 1 & 1 \end{pmatrix} \begin{pmatrix} 1 & 3/5 & \\ & 1 & \\ & & 1 \end{pmatrix} \begin{pmatrix} 1/5 & & \\ & 5 & \\ & & 1 \end{pmatrix} \\
&\qquad \begin{pmatrix} 1 & & \\ 3/5 & 1 & \\ & & 1 \end{pmatrix} \begin{pmatrix} 1 & & 4 \\ & 1 & -1 \\ & & 1 \end{pmatrix} \\
&= \begin{pmatrix} 1 & 3/5 & \\ & 1 & \\ -1 & 2/5 & 1 \end{pmatrix} \begin{pmatrix} 1/5 & & \\ & 5 & \\ & & 1 \end{pmatrix} \begin{pmatrix} 1 & & 4 \\ 3/5 & 1 & 7/5 \\ & & 1 \end{pmatrix}.
\end{aligned}
$$

The final decomposition VDW here is very like an LDU- or UDL-decomposition, but instead of V and W having a triangular pattern, they have what we could call a 'UL/LU' pattern.

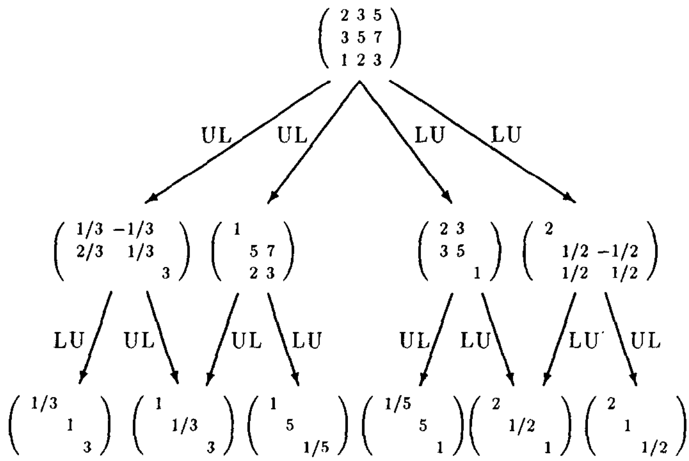


Fig. 2. ULU diagonalizations of a matrix.

Definition 8. A ULU diagonalization, Y of a nondegenerate matrix X is a matrix that can be obtained from X in zero or more block unit UL or LU diagonalization steps.

We say a square nondegenerate matrix X determines a ULU class $[X]$, defined by

$$[X] = \{x \mid x \text{ is an ULU diagonalization of } X\}.$$

Specifically, $X \in [X]$, and the diagonal matrix D in the LDU decomposition of X is also in $[X]$. Fig. 2 shows the entire set $[X]$ of 11 ULU diagonalizations for the matrix X above.

Lemma 3. *Every member of a ULU class* $[X]$ *has determinant* $\det X$.

Proof. LU- and UL-diagonalizations always leave the determinant undisturbed. □

5. ULU classes obey categorial grammar

Both leading and trailing Schur complements define natural quotient operators on ULU classes, and matrix direct sums define a natural product operator.

Formally, let M be the semigroup of matrices under the (associative) multiplicative operation '$\oplus$'. Then three operators can be defined on ULU classes $[X]$, $[Y]$.

$$[X] \oplus [Y] = \{x \oplus y \in M \mid x \in [X], y \in [Y]\}$$
$$[X]/[Y] = \{z \in M \mid \forall y \in [Y], z \oplus y \in [X]\}$$
$$[Y] \setminus [X] = \{z \in M \mid \forall y \in [Y], y \oplus z \in [X]\}.$$

For example, If X is as in the previous example, so $X \trianglerighteq (3)$, then $[X]/[(3)]$ is the class

$$[X]/[(3)] = \left\{ \begin{pmatrix} 1/3 & -1/3 \\ 2/3 & 1/3 \end{pmatrix}, \begin{pmatrix} 1 & \\ & 1/3 \end{pmatrix}, \begin{pmatrix} 1/3 & \\ & 1 \end{pmatrix} \right\}.$$

Notice that $[X]/[(3)] = [X/(3)]$.

Theorem 2. *ULU classes obey the laws of Categorial Grammar. That is: X_1 and X_2 are Categorial Grammar expressions (involving only the three operators above, the empty matrix '1', and variables universally quantified over the space of square matrices) such that $X_1 \Rightarrow X_2$ iff X_1 and X_2 are matrix expressions involving Schur complements and direct sum such that $[X_1] \subseteq [X_2]$.*

Proof. First, ULU classes obey the axioms, where the empty matrix '1' is the identity under direct sum and '$\Rightarrow$' is '$\subseteq$'. The axioms $([X] \oplus 1) = [X]$, $(1 \oplus [X]) = [X]$, and $(([X] \oplus [Y]) \oplus [Z]) = ([X] \oplus ([Y] \oplus [Z]))$ are easy to establish. Second, ULU classes obey the following inference rules as well:

$$[X] \oplus [Y] \subseteq [Z] \quad \text{iff} \quad [X] \subseteq [Z]/[Y].$$
$$[Y] \oplus [X] \subseteq [Z] \quad \text{iff} \quad [X] \subseteq [Y] \setminus [Z].$$

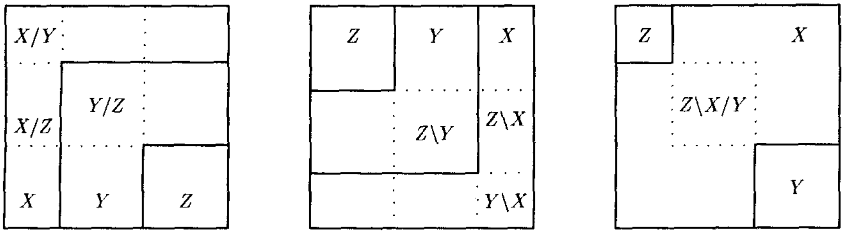
This is a consequence of the fact that ULU classes define the ideals of a ring, and as Lambek notes [17], p. 298, the ideals of a ring always obey the laws of Categorial Grammar. □

A simple pictorial summary of Theorem 2 is possible. Fig. 3 portrays various properties of ULU classes, where Y and Z denote specific block submatrices of X, and '$X_1 \Rightarrow X_2$' means '$[X_1] \subseteq [X_2]$'.

6. Determinants of Schur complements obey categorial grammar

If '$\Rightarrow$' has interpretation 'have equal determinants', Schur complements obey Categorial Grammar.

Definition 9. A matrix expression involving Schur complements and direct sum is well-formed if in every subexpression $(X \ / \ Y)$ both X and Y denote matrices such that $X \trianglerighteq Y$, and in every subexpression $(Y \setminus X)$ both X and Y denote matrices such that $Y \trianglelefteq X$.



$(X/Y) \cdot (Y/Z) \Rightarrow (X/Z)$
$(X/Y) \Rightarrow (X/Z)/(Y/Z)$
$(Y/Z) \Rightarrow (X/Y) \setminus (X/Z)$
$X \trianglerighteq Y \trianglerighteq Z$

$(Z \setminus Y) \cdot (Y \setminus X) \Rightarrow (Z \setminus X)$
$(Y \setminus X) \Rightarrow (Z \setminus Y) \setminus (Z \setminus X)$
$(Z \setminus Y) \Rightarrow (Z \setminus X)/(Y \setminus X)$
$Z \trianglelefteq Y \trianglelefteq X$

$(Z \setminus X)/Y \Rightarrow Z \setminus (X/Y)$
$Z \trianglelefteq X,\ X \trianglerighteq Y$

Fig. 3. Lambek's Categorial Grammar relations, portrayed as properties of ULU classes.

Theorem 3. *Determinants of Schur complements obey the laws of Categorial Grammar. That is: if X_1 and X_2 are well-formed Categorial Grammar expressions such that $X_1 \Rightarrow X_2$ then X_1 and X_2 are matrix expressions involving Schur complements and direct sum such that* det $X_1 =$ det X_2.

Proof. If the matrix expressions X_1 and X_2 are well-formed, then they denote specific matrices A_1 and A_2. We claim that $A_1 \in [X_1]$ and $A_2 \in [X_2]$. This can be proven by inductively establishing (on the size of the expression) each of the following relations:

$$[X \oplus Y] = [X] \oplus [Y],$$
$$[X/Y] \subseteq [X]/[Y],$$
$$[Y \setminus X] \subseteq [Y] \setminus [X].$$

For example in the base case where A is the value of $(X\ /\ Y)$ where X and Y are matrices for which $X \trianglerighteq Y$ (so $(X\ /\ Y)$ is well-formed), then because $[X]/[Y] = \{z \in M \mid \forall y \in [Y], z \oplus y \in [X]\}$ specifically $A \in [X]/[Y]$, and thus $[A] \subseteq [X]/[Y]$, i.e., $[X/Y] \subseteq [X]/[Y]$.

Therefore $[A_1] \subseteq [X_1]$ and $[A_2] \subseteq [X_2]$, so in particular $A_1 \in [X_1]$ and $A_2 \in [X_2]$ as claimed. By Lemma 3, every element of $[X_1]$ (resp. $[X_2]$) shares the common determinant value det $X_1 =$ det A_1 (resp. det $X_2 =$ det A_2). But by Theorem 2, $[X_1] \subseteq [X_2]$, so det $X_1 =$ det X_2. $\square$

7. One-sided Schur complements obey categorial grammar

It is natural to hope Theorem 3 could be strengthened to an 'if and only if' statement, so that det $X_1 =$ det X_2 implies $X_1 \Rightarrow X_2$. Even more ambitious, it would be nice if Categorial Grammar could be related with Schur complement *matrices*, rather than Schur complement *determinants*. This hope is encouraged by the example above, in which $[X]/[(3)] = [X/(3)]$. Moreover, we have identities like the quotient property (Lemma 2) which match quotient reduction laws of Categorial Grammar.

Unfortunately Theorem 3 is no longer valid as an if and only if statement, and is also invalid if we omit the 'det's from its statement. In particular, although it is true that

$$\det\,(\,X\,/\,(\,A\,\setminus\,X\,)) = \det\,A,$$

in general Schur complements do not satisfy the type raising law, i.e.,

$$(\,X\,/\,(\,A\,\setminus\,X\,)) \neq A.$$

For example, with the matrices

$$X = \begin{pmatrix} 1 & 0 & -1 \\ 0 & 1 & 1 \\ 1 & 1 & 1 \end{pmatrix}, \quad A = \begin{pmatrix} 1 & 0 \\ 0 & 1 \end{pmatrix}$$

$$(\,X\,/\,(\,A\,\setminus\,X\,)) = (\,X\,/\,(1)) = \begin{pmatrix} 2 & 1 \\ -1 & 0 \end{pmatrix} \neq A.$$

The basic problem is that a matrix

$$X = \begin{pmatrix} A & B \\ C & D \end{pmatrix}$$

has both the UL-diagonalization $(\,X\,/\,D\,) \oplus D$ and the LU-diagonalization $A \oplus (\,A\,\setminus\,X\,)$, but the inference rules

$$\begin{aligned} &\text{if} \quad (Y \oplus Z) \Rightarrow X \quad \text{then} \quad Y \Rightarrow (X/Z) \\ &\text{if} \quad Y \Rightarrow (X/Z) \quad \text{then} \quad (Y \oplus Z) \Rightarrow X \end{aligned}$$

where '$\Rightarrow$' means 'is a ULU diagonalization of' are valid only for the UL-diagonalization. In particular, although

'$A \oplus (A\,\setminus\,X)$ is a ULU diagonalization of X' is true

generally

'A is a ULU diagonalization of $(X\,/\,(A\,\setminus\,X))$' is false.

However, consider the following 'one-sided' fragment of the Categorial Grammar theory intended for use with Schur complements, where '$\Rightarrow$' means 'is an LU diagonalization of':

Axioms	$(X \oplus 1) \Rightarrow X$	
	$X \Rightarrow (X \oplus 1)$	
	$(1 \oplus X) \Rightarrow X$	
	$X \Rightarrow (1 \oplus X)$	
	$((X \oplus Y) \oplus Z) \Rightarrow (X \oplus (Y \oplus Z))$	
	$(X \oplus (Y \oplus Z)) \Rightarrow ((X \oplus Y) \oplus Z)$	
	$(X/X) \Rightarrow 1$	
	$((X/Y) \oplus Y) \Rightarrow X$ (provided $X \unrhd Y$)	
Rules of	if $X \Rightarrow Y$ and $Y \Rightarrow Z$	then $X \Rightarrow Z$
Inference	if $(Y \oplus Z) \Rightarrow X$	then $Y \Rightarrow (X/Z)$ (provided $X \unrhd Z$)
	if $Y \Rightarrow (X/Z)$	then $(Y \oplus Z) \Rightarrow X$ (provided $X \unrhd Z$)
	if $Y \Rightarrow Z$	then $(X \oplus Y) \Rightarrow (X \oplus Z)$
	if $Y \Rightarrow Z$	then $(Y \oplus X) \Rightarrow (Z \oplus X)$

By design, these rules address only LU-diagonalizations, and cannot produce the confusion above.

Theorem 4. *Schur complements obey the one-sided axioms and inference rules above (a fragment of the Categorial Grammar theory). Specifically: if X_1 and X_2 are expressions such that $X_1 \Rightarrow X_2$ with the axioms and inference rules above then X_1 and X_2 are matrix expressions involving Schur complements and direct sum, and X_1 is an LU diagonalization of X_2. Furthermore, if X_1 does not involve direct sum, then $X_1 = X_2$.*

Proof. The first statement follows since each axiom and inference rule holds of LU diagonalizations. When X_1 does not involve direct sum, it denotes a full matrix (in general). Since X_1 is a diagonalization of X_2, it must then also be the case that X_2 is full, and $X_1 = X_2$. □

With these axioms and inference rules, we can derive the Schur complement quotient law

$$((A\ /\ C)\ /\ (B\ /\ C)) = (A\ /\ B) \quad \text{(provided } A \unrhd B, B \unrhd C\text{)}$$

(Lemma 2) in the following way, without resorting to complex block matrix formulas:

1. $(B/C) \oplus C \Rightarrow B$
   axiom: $X/Y \oplus Y \Rightarrow X$ (provided $B \unrhd C$)
2. $(A/B) \oplus ((B/C) \oplus C) \Rightarrow (A/B) \oplus B$
   1. & rule: if $Y \Rightarrow Z$ then $X \oplus Y \Rightarrow X \oplus Z$
3. $(A/B) \oplus B \Rightarrow A$
   axiom: $X/Y \oplus Y \Rightarrow X$ (provided $A \unrhd B$)

4. $(A/B) \oplus ((B/C) \oplus C) \Rightarrow A$
   2. & 3. & rule: if $X \Rightarrow Y$ and $Y \Rightarrow Z$ then $X \Rightarrow Z$
5. $((A/B) \oplus (B/C)) \oplus C \Rightarrow (A/B) \oplus ((B/C) \oplus C))$
   axiom: $((X \oplus Y) \oplus Z) \Rightarrow (X \oplus (Y \oplus Z))$
6. $((A/B) \oplus (B/C)) \oplus C \Rightarrow A$
   5. & 4. & rule: if $X \Rightarrow Y$ and $Y \Rightarrow Z$ then $X \Rightarrow Z$
7. $((A/B) \oplus (B/C)) \Rightarrow (A/C)$
   6. & rule: if $Y \oplus Z \Rightarrow X$ then $Y \Rightarrow (X/Z)$
8. $(A/B) \Rightarrow ((A/C)/(B/C))$
   7. & rule: if $Y \oplus Z \Rightarrow X$ then $Y \Rightarrow (X/Z)$.

The final consequent

$$(A\ /\ B) \text{ is an LU diagonalization of } ((A\ /\ C)\ /\ (B\ /\ C))$$

actually yields the quotient law equation since the left side does not involve direct sum. Notice also that the seventh step of the derivation above gives the composition law

$$(A/B) \oplus (B/C) \text{ is an LU diagonalization of } (A/C)$$

– a potentially 'new' property of Schur complements.

Many other properties follow from these axioms and inference rules. For example

$$(A \oplus (B\ /\ C)) \text{ is an LU diagonalization of } ((A \oplus B)\ /\ C) \quad (\text{provided } B \trianglerighteq C)$$

is derivable in the following way:

1. $(B/C) \oplus C \Rightarrow B$
   axiom: $X/Y \oplus Y \Rightarrow X$ (provided $B \trianglerighteq C$)
2. $A \oplus ((B/C) \oplus C) \Rightarrow A \oplus B$
   1. & rule: if $Y \Rightarrow Z$ then $X \oplus Y \Rightarrow X \oplus Z$
3. $(A \oplus (B/C)) \oplus C \Rightarrow A \oplus ((B/C) \oplus C)$
   axiom: $((X \oplus Y) \oplus Z) \Rightarrow (X \oplus (Y \oplus Z))$
4. $(A \oplus (B/C)) \oplus C \Rightarrow A \oplus B$
   3. & 2. & rule: if $X \Rightarrow Y$ and $Y \Rightarrow Z$ then $X \Rightarrow Z$
5. $A \oplus (B/C) \Rightarrow (A \oplus B)/C$
   4. & rule: if $Y \oplus Z \Rightarrow X$ then $Y \Rightarrow (X/Z)$.

Studying this property suggests that the rules are not yet complete: the property should yield an equation, yet the converse does not appear to be derivable. Still, the converse holds only provided that $B \trianglerighteq C$.

8. Indexed Schur complements obey categorial grammar

So far we have focused on Schur complements with respect to leading and trailing principal submatrices. This can be generalized to permit arbitrary principal submatrices.

Definition 10. We define a set of indices $\alpha \subseteq \{1, \ldots, n\}$ to denote an *increasing sequence*, and the *complement* $\overline{\alpha}$ of α in $\{1, \ldots, n\}$ to denote the increasing sequence of indices not in α.

For any $n \times n$ matrix X, and any subset α of $\{1, \ldots, n\}$ such that $X[\alpha\,|\,\alpha]$ is nonsingular, the *indexed Schur complement* of $X[\alpha\,|\,\alpha]$ **in** X is

$$(\, X \,/\, X[\alpha\,|\,\alpha]\,) = \Big(\, X[\overline{\alpha}\,|\,\overline{\alpha}] - X[\overline{\alpha}\,|\,\alpha] \cdot (X[\alpha\,|\,\alpha])^{-1} \cdot X[\alpha\,|\,\overline{\alpha}]\Big).$$

Categorial Grammar arises again in this context, because we can treat increasing sequences as sets of indices, and sequence complements as set complements. Let '1' denote the empty set, let α, β, γ be variables denoting arbitrary subsets of a finite set of indices $\mathscr{I} = \{1, \ldots, n\}$, and let the Categorial Grammar operators denote the following operators on sets:

$$\alpha \cdot \beta = \alpha \cup \beta, \quad \gamma/\beta = \gamma - \beta, \quad \alpha \setminus \gamma = \gamma - \alpha.$$

Theorem 5. *Simple index expressions obey Categorial Grammar. Specifically:* α_1 *and* α_2 *are expressions of Categorial Grammar such that* $\alpha_1 \Rightarrow \alpha_2$ *iff* α_1 *and* α_2 *are expressions involving the three operators above that denote subsets of a finite ordered set of indices* $\mathscr{I}$ *such that* $\alpha_1 \supseteq \alpha_2$.

Proof. With the operators having the interpretation above, subsets of indices obey all the axioms, provided that '1' denotes the empty set. They also obey the rules of inference, since they satisfy the properties

$$\alpha \cdot \beta \supseteq \gamma \quad \text{iff} \quad \alpha \supseteq \gamma/\beta.$$
$$\alpha \cdot \beta \supseteq \gamma \quad \text{iff} \quad \beta \supseteq \alpha \setminus \gamma.$$

They thus give a model of Categorial Grammar. □

By extending the argument above so that we write $X[\alpha\,|\,\alpha]$ instead of α, and change the operator denotations accordingly, we find indexed Schur complements also obey Categorial Grammar.

Lemma 4 (Schur's identity for determinants). *If* α *and* β *are disjoint, then*

$$\frac{\det\, X[\alpha \cdot \beta\,|\,\alpha \cdot \beta]}{\det\, X[\beta\,|\,\beta]} = \det\, (X[\alpha \cdot \beta\,|\,\alpha \cdot \beta] \,/\, X[\beta\,|\,\beta]).$$

Proof. This is derivable from LU decomposition of $X[\alpha \cdot \beta \mid \alpha \cdot \beta]$, which finds the diagonal entries of U to be $X[\alpha \mid \alpha]$ and $(X[\beta \mid \beta] - X[\beta \mid \alpha] (X[\alpha \mid \alpha])^{-1} X[\alpha \mid \beta])$. □

Brualdi and Schneider [4] have characterized a wide class of determinantal identities as a prime ideal generated by an intuitive set of elementary identities. It would be interesting if Theorem 5 and Lemma 4 could be extended to something like this, asserting that Categorial Grammar gives a complete theory for determinantal identities involving indexed Schur complements, i.e., produces all possible determinantal identities.

An alternative to indexed Schur complements is to retain the unindexed Schur complements defined earlier, but allow identical permutation of rows and columns of X, so that the matrix can be reordered to place any desired principal submatrix in the upper left corner. Some authors have proposed incorporating permutation into Categorial Grammar, motivated by better modeling of natural language. Sometimes permutation is incorporated by adding the equivalence

$$(X \,/\, Y) \equiv (Y \setminus X),$$

but refinements of this simple approach have also been proposed [14,18].

9. Concluding remarks

This paper has studied several natural connections between Schur complements and Categorial Grammar, and shown how Categorial Grammar can lend perspective on LU decomposition and Schur complements. For example, Fig. 3 suggests identities satisfied by Schur complements, and Categorial Grammar makes it possible to derive the quotient property of Schur complements without complex matrix manipulations. It should be possible to expand this connection in several ways, perhaps including the following.

9.1. Matrix and Matroid theory

The results above hint at more fundamental connections between general Categorial Grammar and Schur complements. In particular:

Duals, Adjoints, and Inverses: Reversal has an important role in formal grammar, and we can introduce a dual operator on matrices defined by $X^R = RXR$ where $R = (r_{i,j})$ is the reversal permutation defined by $r_{i,j} = 1$ if $i = (n - j + 1)$, 0 otherwise. Then

$$(X \oplus Y)^R = (Y^R \oplus X^R),$$
$$(X / Y)^R = (Y^R \setminus X^R),$$
$$(X \setminus Y)^R = (Y^R / X^R).$$

Also there are several matrix adjoints '$\star$' that have the following slightly different properties:

$$(X \oplus Y)^\star = (X^\star \oplus Y^\star),$$

$$(X / Y)^\star = (X^\star / Y^\star),$$

$$(X \setminus Y)^\star = (X^\star \setminus Y^\star).$$

For example, these identities hold when '$\star$' is either matrix transpose or Hermitian adjoint (conjugate transpose). Finally various properties also hold for matrix inverse:

$$(X \oplus Y)^{-1} = (X^{-1} \oplus Y^{-1}),$$
$$(X / Y)^{-1} \setminus X^{-1} = Y^{-1},$$
$$X^{-1} / (Y \setminus X)^{-1} = Y^{-1}.$$

Recent work in Lambek systems has considered the use of unary operators. For example, Moortgat [19] introduces modal operators with the following new rules:

if $\diamond X \Rightarrow Y$ then $X \Rightarrow \Box Y$,
if $X \Rightarrow \Box Y$ then $\diamond X \Rightarrow Y$

In addition, he considers the effects of including a variety of modal axioms, including among others the distributivity postulates

$$\Box(X/Y) \Rightarrow (\Box X)/(\Box Y),$$
$$\diamond(X \cdot Y) \Rightarrow (\diamond X) \cdot (\diamond Y),$$

resembling the adjoint identities above.

Minors and Matroids: Generalizations of both the preceding ideas have been studied in the framework of matroids. Specifically, Oxley ([23], ch. 3) reviews related results of Tutte [28] on 'minors' of matroids. Any matrix X defines a matroid M of linearly independent sets of columns of X, and the complements of these sets define the dual matroid M^*. If T is a set of columns of X, the natural quotients satisfy the identity

$$M \;/\; T = (M^* \;\setminus\; T)^*,$$

and minors of the matroids can be defined that capitalize on this duality.

9.2. Gaussian elimination practice

The Schur complement has gradually gained recognition as a useful tool in various aspects of matrix analysis. The results collected in this paper help summarize why, and suggest it may be interesting to have Gaussian elimination generate ULU decompositions, rather than just LU decompositions. That is, let a UL/LU (ULU) decomposition $X = VYW$ of a matrix X give a diagonal matrix Y and a pair of matrices V and W resulting from recursive block unit LU- or UL-diagonalization. The increased flexibility of being able to select either LU or UL decomposition dynamically could conceivably permit fast matrix decomposition with better roundoff properties than ordinary LU decomposition.

Acknowledgements

I am indebted to Michael Moortgat for pointing out an error in an earlier version of this paper, for giving me an overview of some recent work on Categorial Grammar, and for informing me about generalizations of Categorial Grammar involving unary operators. I am also indebted to an anonymous referee, who provided excellent suggestions on how to improve the paper.

References

[1] K. Ajdukiewicz, Die Syntaktische Konnexität, Studia Philosophica, 1 (1935) 1–27.
[2] Y. Bar-Hillel, Language and Information: Selected Essays on Their Theory and Application, Addison–Wesley, Reading, MA, 1964.
[3] J. van Benthem, The Lambek Calculus, in: R.T. Oehrle, E. Bach, D. Wheeler (Eds.), Categorial Grammars and Natural Language Structures, Reidel, Boston, 1988, pp. 35–68.
[4] R. Brualdi, H. Schneider, Determinantal identities: Gauss, Schur, Cauchy, Sylvester, Kronecker, Jacobi, Binet, Laplace, Muir, Cayley,, Linear Algebra Appl. 53 (1983) 769–791.
[5] D. Carlson, E. Haynsworth, T. Markham, A generalization of the Schur complement by means of the Moore-Penrose inverse, SIAM J. Appl. Math. 8 (1) (1974) 169–175.
[6] R.W. Cottle, Manifestations of the Schur complement, Linear Algebra Appl. 8 (1974) 189–211.
[7] D.E. Crabtree, E.V. Haynsworth, An identity for the Schur complement of a matrix, Proc. AMS 22 (1969) 364–366.
[8] T. Evans, On multiplicative systems defined by generators and relations, Proc. Cambridge. Philos. Soc. 47 (1951) 637–649.
[9] F.R. Gantmacher, The Theory of Matrices, vol. 1, Chelsea, New York, 1959.
[10] E. Haynsworth, Determination of the inertia of a partitioned Hermitian matrix, Linear Algebra and Appl. 1 (1968) 73–81.
[11] E. Haynsworth, Applications of an inequality for the Schur complement, Proc. Amer. Math. Soc. 24 (1970) 512–516.
[12] E. Haynsworth, Reduction of a matrix using properties of the Schur complement. Linear Algebra and Appl. 3 (1970) 23–29.
[13] R.A. Horn, C.R. Johnson, Matrix Analysis, Cambridge University Press, New York, 1992.
[14] E.L. Keenan, A. Timberlake, Natural Language Motivations for Extending Categorial Grammar, in: R.T. Oehrle, E. Bach, D. Wheeler (Eds.), Categorial Grammars and Natural Language Structures, Reidel, Boston, 1988, pp. 265–295.

[15] J. Lambek, The Mathematics of Sentence Structure, Amer. Math. Monthly 65 (1) (1958) 154–170.
[16] J. Lambek, Rings and Modules, Chelsea, New York, 1966.
[17] J. Lambek, Categorial and Categorical Grammars, in: R.T. Oehrle, E. Bach, D. Wheeler (Eds.), Categorial Grammars and Natural Language Structures, Reidel, Boston, 1988, pp. 297–317.
[18] M. Moortgat, Categorial Investigations: Logical and Linguistic Aspects of the Lambek Calculus, Foris Publications, Providence, RI, 1988.
[19] M. Moortgat, Multimodal linguistic inference, in: Gabbay, Kempson (Eds.), Language and Proof Theory, special issue of Journal of Logic, Language and Inference (JoLLI), 1996.
[20] M. Moortgat, Categorial type logics, in: J. van Benthem, A. ter Meulen (Eds.), Handbook of Logic and Language, Elsevier, Amsterdam, 1996.
[21] R.T. Oehrle, E. Bach, D. Wheeler (Eds.), Categorial Grammars and Natural Language Structures, Reidel, Boston, 1988.
[22] A. Ostrowski, "A new proof of Haynsworth's quotient formula for Schur complements", Linear Algebra Appl. 4 (1971) 389–392.
[23] J.G. Oxley, Matroid Theory, Oxford University Press, New York, 1992.
[24] D.S. Parker, Random butterfly transformations with applications in computational linear algebra, Technical Report CSD-950023, UCLA Computer Science Department (1995).
[25] D.S. Parker, Explicit formulas for the results of Gaussian elimination, Technical Report CSD-950025, UCLA Computer Science Department (1995).
[26] D.S. Parker, D. Le, How to eliminate pivoting from Gaussian elimination – by randomizing instead, Technical Report CSD-950022, UCLA Computer Science Department (1995).
[27] I. Schur, Über Potenzreihen die im Innern des Einheitskreises beschränkt sind, J. Reine Angew. Math. 147 (1917) 205–232.
[28] W.T. Tutte, A Homotopy theorem for matroids. I, II, Trans. AMS 88 (1958) 144–174.